ESD-TR-68-121

MTR-575

SAMPLING THEORY AND THE TESTING OF COMMON-USER COMMUNICATIONS SYSTEMS

MAY 1968

E. H. Harwood

Prepared for

DEPUTY FOR COMMUNICATIONS SYSTEMS
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
L. G. Hanscom Field, Bedford, Massachusetts

This document has been approved for public release and sale; its distribution is unlimited.

Project 4900
Prepared by
THE MITRE CORPORATION
Bedford, Massachusetts
Contract AF19(628)-5165

ESD-TR-68-121

MTR-575

SAMPLING THEORY AND THE TESTING OF COMMON-USER COMMUNICATIONS SYSTEMS

MAY 1968

E. H. Harwood

Prepared for

DEPUTY FOR COMMUNICATIONS SYSTEMS
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
L. G. Hanscom Field, Bedford, Massachusetts

Project 4900
Prepared by

THE MITRE CORPORATION
Bedford, Massachusetts

Contract AF19(628)-5165

FOREWORD

The work reported in this document was performed by The MITRE Corporation, Bedford, Massachusetts for the Deputy For Communications Systems, Electronic Systems Division, of the Air Force Systems Command under Contract AF 19(628)-5165.

REVIEW AND APPROVAL

Publication of this technical report does not constitute Air Force approval of the report's findings or conclusions. It is published only for the exchange and stimulation of ideas.

SHERMAN R. CUMMINGS, Colonel, USAF
System Program Director
Long Lines Communications SPO
Deputy for Communications Systems

ABSTRACT

The techniques of sampling theory are applied to measure the performance of the plant of a common-user communication system.

Based on a range of expected values of p (probability of call failure), appropriate values are derived for n (the sample size), and the rationale behind the selection of a confidence level is explained. Where sampling results indicate an identifiable cause of ineffective calls, remedial engineering work can be undertaken. By this process system performance can be improved until no readily identifiable cause of failure can be found.

The work in this report was done in order to provide input to Annex H of the Overseas AUTOVON Master Test Plan.

ACKNOWLEDGMENT

The author's best thanks are due Bruce A. Van Auken for his guidance in preparing these notes.

TABLE OF CONTENTS

SECTION I

INTRODUCTION

"Statistics is the art of stating in precise terms that which one does not know."

--Statistician's epigram

The "common-user" communications system under consideration in this report consists of one or more switching centers to each of which many subscribers are connected by access lines. The switching centers are interconnected by groups of interswitch trunks. The subscribers originate telephone calls to each other at times of their own choosing. The performance of the system can be specified in terms of the number of calls that fail.

One source of ineffective calls is the failure of plant (equipment and/or circuits) to operate in the manner for which it was designed; i.e., some item of plant involved in processing the call is faulty. This report deals with the performance of the common-user system when call attempts fail because of such plant malfunction.

In order to eliminate the irrelevant complications arising from busy conditions, the rule will be imposed that the probability of encountering a busy condition in the system must be very small compared with the probability of ineffective calls due to faulty plant. This is equivalent to saying that the system must be lightly loaded.

When the performance of the plant of a system is specified in terms of the number of calls that fail because of faulty plant (equipment and/or circuits), a system "figure of merit" readily follows. Given such a measure, the performance of the plant in different systems can be compared, one with

another, and criteria of goodness or badness can be established and used. One such measure, termed Plant Performance Index (PPI), is the fraction obtained by dividing the number of calls that fail by the total number of call attempts. This fraction expressed as a decimal is similar, qualitatively, to the commonly used "grade of service," e.g., 0.02. It can also be thought of as a proportion. In statistical terms it is the failure probability, and this term will be used throughout this report.

The failure probability is likely to be greater in a newly structured system – such as one containing switching centers of a new design, for example, or one using transmission circuits not originally designed as a part of a switched system (or both) – than in an integrated system, i.e., one that has "shaken down" and where any frequently recurring design or manufacturing deficiencies have been rectified. In an integrated system, independent failures will still occur in any of the various individual components that make up the system, but no one single type or piece of equipment can be singled out as a recurring prime cause of failure.

As the number of call attempts increases, the proportion that fail will tend towards a stable value. The larger the number of attempts the narrower becomes the range of fluctuation of the proportion around the stable value. Also, as the number of call attempts increases, so does the likelihood that the many diverse potential causes of failure will manifest themselves. The net result is that the behavior of the system plant can be characterized by a failure probability.

Ideally, therefore, we are led to accept the necessity of making an infinitely large number of calls and counting those that fail. In the practical world this is clearly not possible, so this failure probability is inherently a conceptual device. Fortunately, however, sampling theory provides us with

techniques by means of which the concept can be given practical significance. The engineer who is faced with the problem of evaluating this probability in real systems is able to draw useful numerical inferences about it from the results of relatively few call attempts by using the theory of the sampling process.

In this report the symbol "p" will mean this call failure probability in the infinite population of call attempts, and 1-p therefore indicates the corresponding probability of success.

SECTION II

THEORY OF THE SAMPLING PROCESS

For any given size of sample of call attempts, say n in number, the possible number of ineffective calls ranges all the way from zero to n in discrete steps of one. Each of these possible call failure numbers has a certain probability of occurrence; in mathematical terms, the probability that any particular number x of ineffective calls will result is

$$P_{(x)} = \frac{n!}{(n - x)!\,x!} p^{x} (1 - p)^{n - x} .$$

Thus, the probability that no ineffective calls will occur is $(1 - p)^n$ and the probability that all calls will be ineffective is p^n. (When checking these, remember that $0!$, p^0 and $(1 - p)^0$ are all equal to 1).

Within the sample of n call attempts there is thus an associated distribution of probabilities of ineffective calls, and this is called a sampling distribution. An example of such a distribution is given in Figure 1* for the case when $n = 300$ and $p = 0.02$. These values were chosen as being of the right order of magnitude for some practical cases. In the figure, the abscissa Scale A shows the number of ineffective calls (NIC) in the sample, and Scale B shows the proportion of ineffective calls (PIC) in the sample. PIC = NIC/n.

The curve is incomplete as shown – it actually extends far to the right with Scale A reaching 300 and Scale B 1.0. In these extreme-right regions

---

*Data for Figure 1 were obtained from Tables of the Cumulative Binomial Probability Distribution. [1]

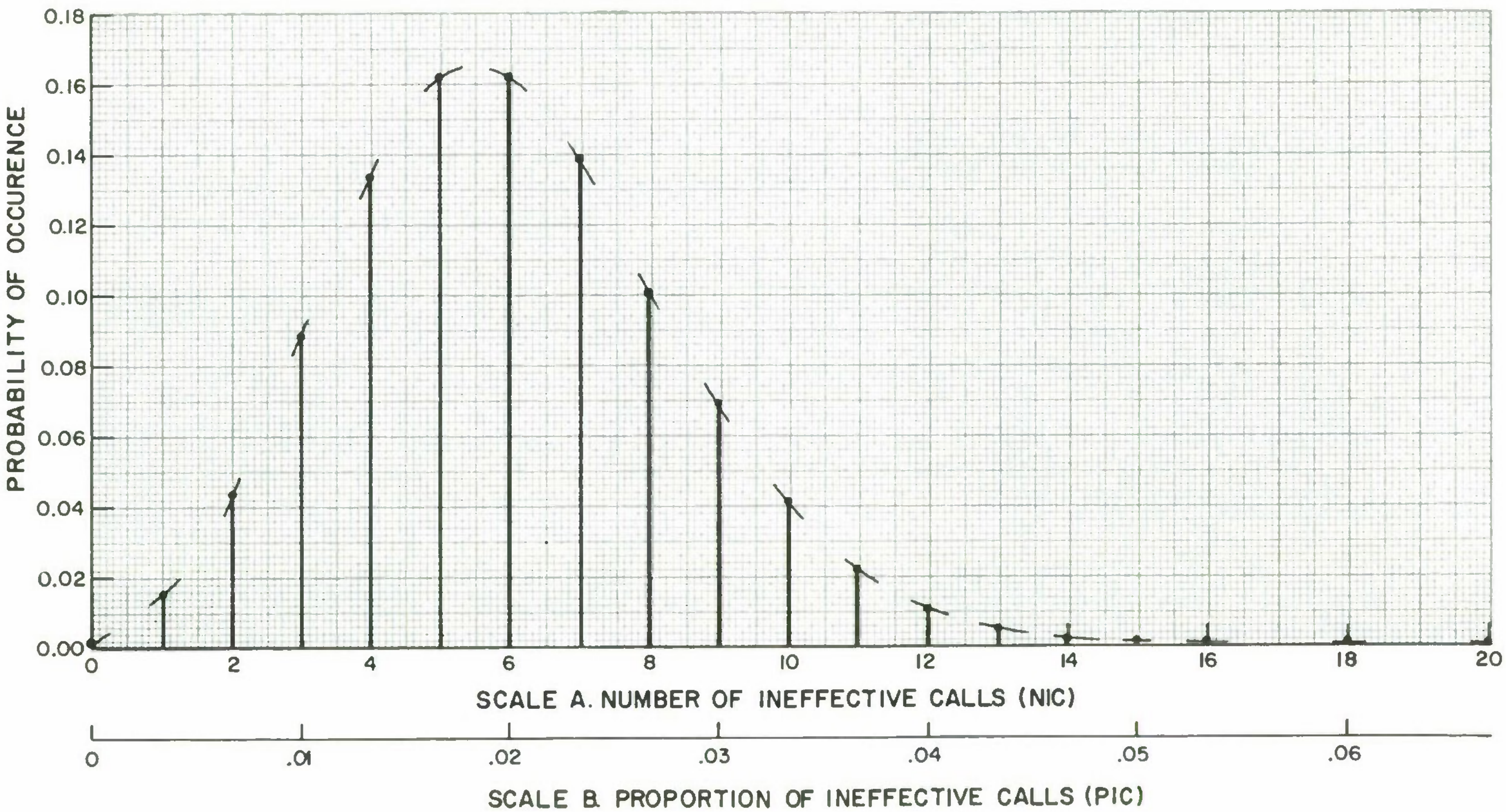

Figure 1. Probability Distribution of Sample Proportions and Number of Ineffective Calls

the probabilities are insignificantly small; i.e., such values of x are "most improbable." It should be noted that the "most probable" value for x is when NIC = 6, but the associated probability of occurrence of this "most probable" value is still quite low at 0.16.

The shape of the curve is very similar to the bell shape of the normal distribution, thus illustrating an important general law of statistics that "measures computed from (random) samples usually tend to be normally distributed." This law, the Central Limit Theorem, coupled with the properties of normal distributions, permits us to use the results of sampling processes to make inferences about the properties of the population from which the samples were taken. When these general remarks are applied specifically to the situation shown in Figure 1, they mean that, given the number of ineffective calls in the samples, each consisting of 300 call attempts, it is possible to state with a certain degree of confidence (expressed as a percentage) that the value of p lies within certain limits. For the present problem it can safely be assumed that the shape of the distribution of possible ineffective calls will be normal if the size of the sample is greater than, say, 100. (For other reasons, also, as will be seen later, the sample size will be kept in the low hundreds.)

Any normal distribution is completely defined when its mean and standard deviation are known. The especial significance of the standard deviation in our problem is that it indicates the degree to which the possible values of the variable are clustered or grouped about the mean – the greater the value of the standard deviation the wider the spread, and vice versa. When the standard deviation is small, a large percentage of the possible values of the variable will be found close to the mean.

We recognize only two outcomes from call attempts – they succeed or they fail. It is a binary situation, not an analog one. In such cases, the mean value of the sampling distribution is equal to np when the abscissa represents number of ineffective calls (Scale A of Figure 1), or is equal to p if the proportion of ineffective calls is plotted (Scale B of Figure 1). For such "binary" populations statisticians use the term "standard error" in place of "standard deviation," and this is equal to $\sqrt{\frac{p(1 - p)}{n}}$. Figure 2 shows how the standard error varies with sample size n for different values of p. It will be seen that for the lower values of n the standard error drops rapidly as n increases, but that when n is about 300 or more this rate of change in the standard error slows down considerably.

Based on these theoretical considerations, it is possible, as has already been suggested, to make some positive assertions regarding the outcome of the sampling process. For example, if sampling is performed under certain conditions (to be discussed later) then the results of the sampling will produce a distribution of the normal kind. The statistician is, therefore, "confident" of the outcome of the sampling process and expresses his degree of confidence in percentages. Since in any normal distribution about 68 percent of all possible outcomes will be within ±1 standard error of the mean of the distribution, he therefore rates his confidence coefficient as 68 percent that any one outcome will be found within the same confidence interval.

In the particular situation under study, the mean of the sampling distribution (Scale B) is p. Hence, we can be 68 percent confident that the true value of p (unknown to us) will lie within an interval bounded by ±1 standard error from the PIC value of the sample.

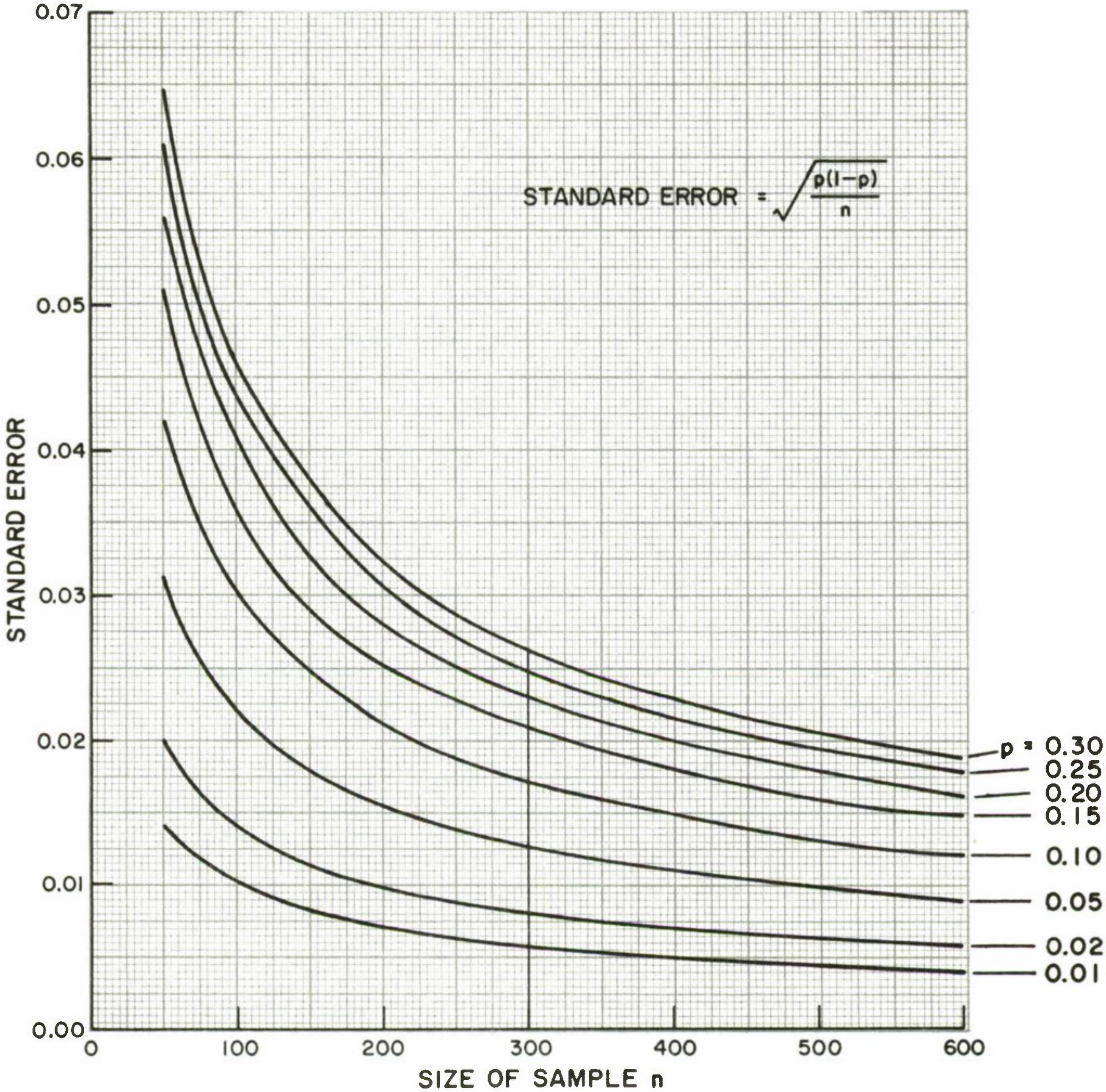

Figure 2. Variation of Standard Error with Sample Size

Any required confidence coefficient other than 68 percent can be obtained from tables of the normal distribution. The table gives some useful approximate values:

Table I

Number of Standard Errors	Confidence Coefficient (%)
0.68	50
0.84	60
1.00	68
1.04	70
1.28	80
1.65	90
2.00	95.5
2.58	99

As was suggested earlier, certain conditions must be observed in our sampling in order not to invalidate the theory. These are:

(a) That the outcome of each call attempt is independent of (i.e., not influenced by) all other call attempts.

(b) That the call attempts making up the samples are truly random choices, i.e., specified by a deliberate process which gives the same probability of being selected to each and every possible combination of calling and called subscribers. This is more difficult than it sounds at first sight and is most easily accomplished in practice by using tables of random numbers.

SECTION III

PREPARATION FOR PRACTICAL APPLICATION

It is manifestly desirable to try to achieve a high confidence coefficient and, at the same time, a small standard error, inasmuch as such a combination reduces the uncertainty in the estimate of p to a minimum. Figure 2 shows, however, that the values of standard error become smaller only as the sample size n increases, and that for any particular value of n the standard error increases with increase in p. Both of these effects are contained in the formula,

$$\text{Standard Error} = \sqrt{\frac{p(1 - p)}{n}} .$$

In preparation for the practical application of sampling theory, the three basic questions then are:

1. What values of p may be expected in working systems?
2. What size should n be?
3. What percentage confidence should be used?

With regard to the first question the author, in discussing the plant performance of AUTOVON communications (mentioned earlier) has provided reasoning which suggests that values of p ranging approximately from, say, 0.02 up to perhaps 0.30 might be met with in practice. The lower end of the range would, it is suggested, apply to the less complex subsystems consisting of one switching center and its dependent subscribers, while the upper end of the range might be found in multi-switch systems with global coverage. Accordingly, values of p ranging from 0.01 to 0.30 were used in calculating the standard error curves of Figure 2.

In considering the size of n the need for standard error values as small as possible has already been established, but unfortunately such can be attained only for the higher values of n. The practical difficulties of sampling – the time needed for a program of test call samples, the central organization required, and the necessity for a number of reasonably skilled testers coupled with the requirement that the system under test be lightly loaded – all these factors mitigate against the use of high values for n. Also, it has to be recognized that the practical difficulties of sampling will probably increase when higher and higher values of p are encountered, as anticipated in systems that will cover geographical areas of continental dimensions. When such considerations are coupled with the trends of the curves depicted in Figure 2, values of n around 300 to 400 suggest themselves as reasonable compromises. In Annex H of the Overseas AUTOVON Master Test Plan, where the author has used these techniques, 300 was specified as a practical value of n. However, if initial tests taken with n = 300 indicate that p is likely to be found amongst the lower values used for Figure 2, then this knowledge will justify a reduction in sample size to 200, or even 100, if further testing is necessary.

The final factor to be appreciated is the shape of the curve linking the percentage confidence with the corresponding number of standard errors. Figure 3 shows the table from page 9 in graphic form. A reasonably linear relationship exists between the two factors up to 85 percent values of confidence. Higher values of confidence are attained only at the expense of a rapidly increasing confidence interval. An 85 percent confidence coefficient is recommended for use. The corresponding confidence interval is the PIC value ± S (where S, the spread on either side of the mean, equals 1.44 times the standard error).

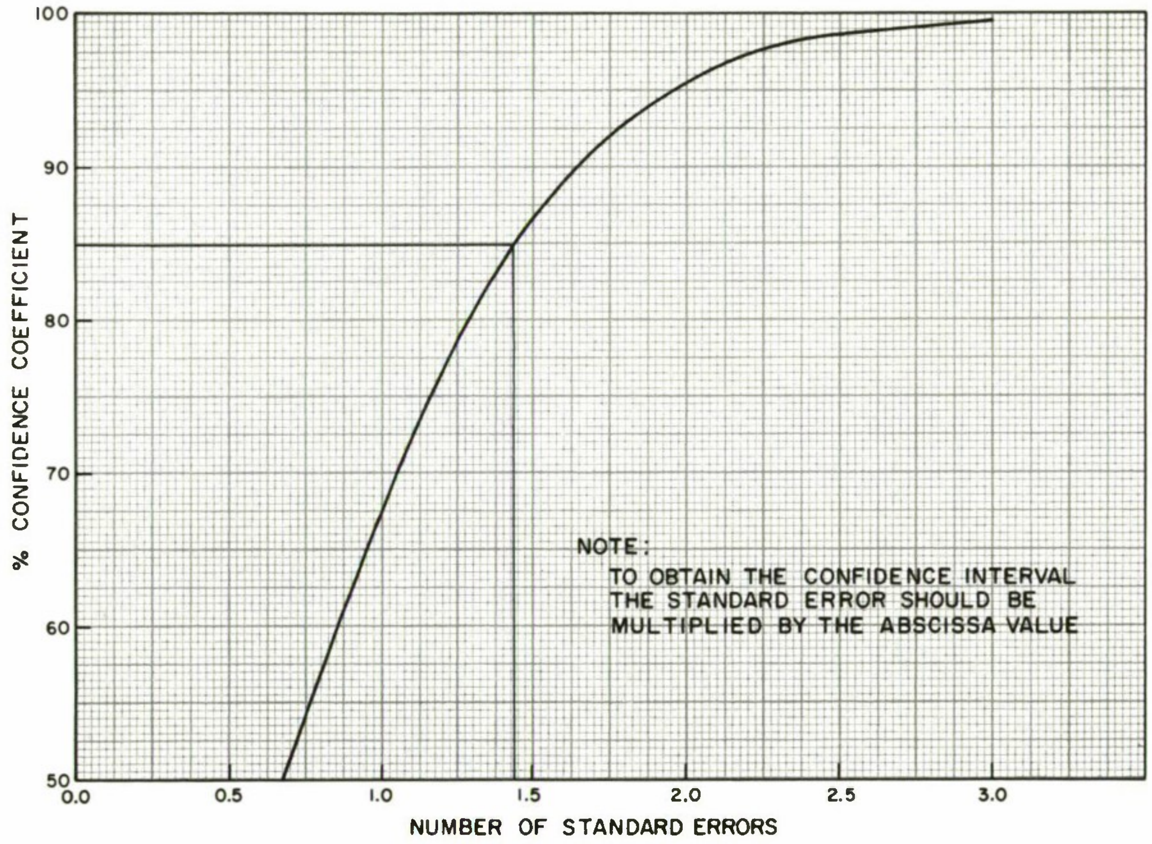

Figure 3. Confidence Coefficient and Number of Standard Errors

Figure 4 gives the variation of S with PIC for the ease when $n = 300$ and the confidenee eoefficient is 85 pereent. The confidenee interval within whieh p will lie is then PIC ± S.

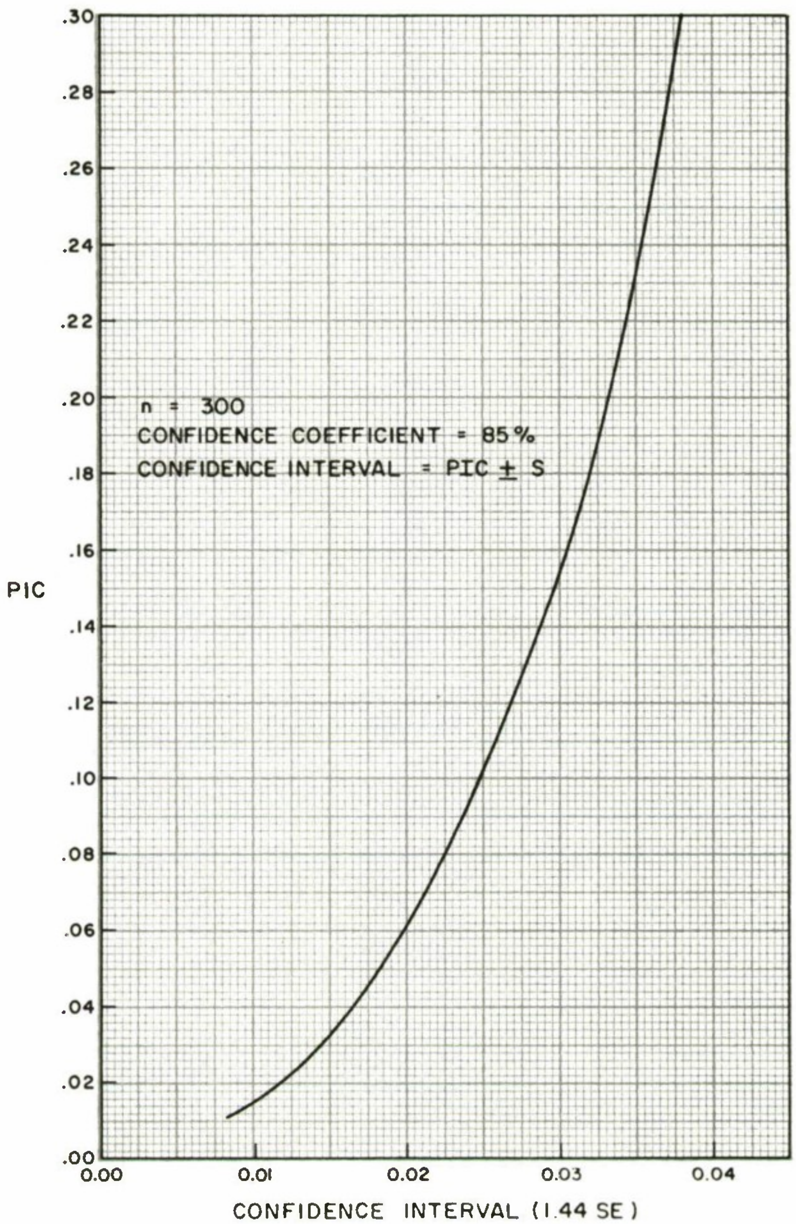

Figure 4. Derivation of Confidence Interval

SECTION IV

PRACTICAL APPLICATION

In the preceding section recommendations have been made and supported regarding appropriate choices for n and the confidence percentage. These choices were made on the assumption that, in practice, the range of possible values for p might lie somewhere between 0.01 and 0.30, according to the complexity (or quantity) of plant involved and the quality of its performance.

After a newly structured system has been tested in the field by measuring the number of ineffective calls in a sample size n , the PIC must be determined (PIC = NIC/n). Next, the resulting PIC value is applied to the ordinate scale of Figure 4 and the corresponding value of S is read off from the abscissa; if the PIC value is 0.16, for example, S is approximately 0.031. Hence, it can be said with 85 percent confidence that the true value of p of the system lies somewhere in the range 0.16 ± 0.031, i.e., between 0.129 and 0.191. This is the best that can be done with one sample, size 300. If the opportunity exists to take a second sample of the same size, without any changes to the plant that might affect its performance, then the confidence interval would shrink from 0.031 to 0.031/1.41, i.e., to 0.022, in which case the 85 percent confidence range would extend from 0.138 to 0.182. When the results of the first sampling indicate that an identifiable cause of ineffective calls exists, however, then good engineering practice will demand that this cause be removed by modifications to equipment and/or circuits. After this remedial work has been completed, a fresh sample should be taken, and the improvement in the system performance will be reflected in the lower range of values for p. Ideally, this process should continue until

no readily identifiable causes of failure can be found -- in other words, until the ineffective calls are due to chance effects.

Such expenditure of effort is completely justified, in the author's opinion, when a new and important system element -- a new switching center, for example -- is being field tested for the first time. The improvement in system performance resulting from such work will be reflected in lower p's for all subsequent similar installations, with comprehensive and permanent benefits that should not lightly be foregone.

EHH:dmp

REFERENCES

1. Tables of the Cumulative Binomial Probability Distribution, Cambridge, Massachusetts: Harvard University Press, 1955.

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) The MITRE Corporation Bedford, Massachusetts	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP N/A

3. REPORT TITLE

SAMPLING THEORY AND THE TESTING OF COMMON-USER COMMUNICATIONS SYSTEMS

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

N/A

5. AUTHOR(S) (First name, middle initial, last name)

Edward H. Harwood

6. REPORT DATE May 1968	7a. TOTAL NO. OF PAGES 22	7b. NO. OF REFS 1

8a. CONTRACT OR GRANT NO. AF 19(628)-5165 b. PROJECT NO. 4900 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) ESD-TR-68-121 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) MTR-575

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES N/A	12. SPONSORING MILITARY ACTIVITY Deputy For Communications Systems, Electronic Systems Division, L. G. Hanscom Field, Bedford, Massachusetts.

13. ABSTRACT

The techniques of sampling theory are applied to measure the performance of the plant of a common-user communication system.

Based on a range of expected values of p (probability of call failure), appropriate values are derived for n (the sample size), and the rationale behind the selection of a confidence level is explained. Where sampling results indicate an identifiable cause of ineffective calls, remedial engineering work can be undertaken. By this process system performance can be improved until no readily identifiable cause of failure can be found.

The work in this report was done in order to provide input to Annex H of the Overseas AUTOVON Master Test Plan.

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
SAMPLING THEORY						
COMMON-USER COMMUNICATIONS SYSTEMS						
COMMUNICATIONS SYSTEMS TESTING						